Studies in Avian Biology No. 6:327–333, 1981.

LIMITATION AND VARIABILITY IN HEARING ABILITY IN CENSUSING BIRDS

ANDRÉ CYR[1]

ABSTRACT.—Too few studies have dealt with the human observer's effect on census results. Factors limiting hearing include the physico-acoustical properties of the ear itself. The frequency levels heard, although they cover a large part of the range emitted by birds, are not perceived identically over their entire range. Age decreases the perception of high frequencies. Time interval, resolution of sound, binaurality, sound shadow effect, fatigue, and masking might all impair our perception to a higher degree than is usually believed and thus affect identification and the census results.

Heretofore, studies of variability in census taking have focused primarily on comparisons between observers without reference to a known bird population or known perceptible fraction of it. An experiment designed to compare the efficiency of observers to a known check sample tape recording shows that even audible sounds are easily overlooked, due either to lack of familiarity with a particular song by some observers or to the masking effect of simultaneous songs, or other factors. Some research topics are proposed to improve quality of hearing and efficiency in interpreting of bird songs to gain new insight into the observer's effect on census results.

Although hearing plays an important role in the life of birds (Hinde 1969, Thielcke 1976), it is also important to those who count them. Observers censusing terrestrial birds often spend 75% or more of their census time listening in order to localize or identify birds. The question here is: to what extent is it possible to use hearing ability and still be confident in our census results?

A census taker in the field faces many stimuli, emitted more or less simultaneously, and attempts to differentiate all these stimuli. We try to intercept messages sent primarily to other birds of the same or different species or to other animals, in addition to the information sent to us as potential predators. We try to intercept the information and correctly decode it for censusing purposes. Are our tools adequate to analyze and decode this information properly? To what extent do we succeed in doing so? What can we do to improve our success?

In the following, I will focus on the factors affecting the hearing variability of the observers. After reviewing the sparse literature involving comparisons between results obtained by different observers, I will suggest experiments that should be done to enhance our hearing ability and our knowledge of its drawbacks on potential results, and will report the results of a small number of such experiments.

LIMITATIONS OF HUMAN HEARING

Problems and limitations include the physical nature of the ear, the threshold of audibility, frequency discrimination, the sound shadow effect, fatigue, the masking effect, and environmental noises. Human ears (Burns 1973, Howard 1973) and those of birds (Pumphrey 1961, Schwartzkopff 1973) are anatomically different, but functionally about as efficient. Birds usually produce sound between 500 and 5000 Hz. Human ability to detect pure tones ranges from 16 to 20,000 Hz. We would conclude that we can perceive sounds produced by birds over almost all their frequency range, except for a few extreme species such as the Oilbird (*Steatornis caripensis*) and some other partly echo-locating species. But our ear remains more efficient between 2 and 6 kHz. Audiological measurements usually refer to pure tones, but we very seldom census these in the field. The limitations of our ear could be much greater than is usually believed, either from audiological measurements or census results. Thus, the first improvement depends on the inherent properties of the physical ear, the quality of which can be improved by practice: the more the tool is used, the more efficient it will be.

Threshold of audibility usually varies from person to person and even from day to day and hour to hour (Beranek 1954). After exposure to even a moderate noise level, slight temporary deafness occurs, which shifts the detection threshold upward, but age is the main factor affecting the threshold of audibility. As seen from Burns (1973:102), higher frequencies are lost faster with age than lower ones. Although sound localization can still be achieved by a single ear with fair accuracy, using intensity cues (Howard 1973), threshold of audibility could have a serious bearing on hearing efficiency for census taking. According to the goal and methodology of censusing, these problems might limit the participation in a particular program.

As compared to the human ear, that of a bird is capable of better resolution of sounds emitted at short time intervals, and birds can react to them accordingly (Pumphrey 1961, Knudson 1978). This is best exemplified by the duetting

[1] Département de Biologie, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1.

in such birds as the Little Grebe (*Podiceps ruficollis*) (Thielcke and Blume 1973, Thielcke 1976). Hirsh (1959) reports that two brief sounds will be perceived as separate with only a few msec between them, but it will take intervals of up to 15–20 msec for the listener to report which of the two preceded the other. He says further that this result is independent of the nature of the sound, whether short or long or of high or low frequency. Henning (1966) found further that differences of 300 Hz are necessary for two sounds of high frequencies (10 kHz) to be discriminated correctly at a level of 75%. This time interval component is not of the utmost importance, for we do not need to react to single notes within a song, but it surely could reduce our faculty of song discrimination and might explain our reduced ability to discriminate between structurally comparable songs such as trills. Our ability to interpret census results correctly would be improved by new knowledge on the birds themselves: to what extent does a bird vary its singing within its range or between morphologically and structurally different habitats or within different bird communities?

Another problem lies in the sound shadow effect (Howard 1973). A sound reaching one ear laterally reaches the other slightly later, producing binaural cues that can be used to estimate the distance from which a sound is emitted. Casseday and Neff (1973) found that man uses different cues to localize pure tones of high and low frequency. Around 3–4 kHz, localization is more efficient than at lower or higher frequencies. At higher frequencies, intensity is used as the cue, whereas a time cue is used at lower frequencies, because of the relatively longer time lag of such a sound travelling from one ear to the other. Methodologies taking distances into account should perhaps avoid including species with high frequency utterances.

Fatigue may impair our hearing ability and is one source of systematic variability in that ability. Thus when designing our field experiments, the complexity of their application should be considered in view of this limiting factor (see Ramsey and Scott 1981a).

Masking is defined as the amount by which the threshold of detection of a sound is raised by the presence of another sound, the masker (Studebacker 1973). Fortunately, for census purposes, exposure to low frequency does not affect the threshold of detectability of high frequency sounds (Ward 1966). The contrary holds as well. Although not specifically studied in relation to census taking, some other factors that play a role in bird communication can surely affect our efficiency at locating and identifying some species. For example, Witkin (1977) showed that the directionality of the source as related to the receiver influences the receiver's ease of locating the source of bird communication. In censusing, the receiver is the observer but the problem remains, although little attention has been paid to that point (see also Wiley and Richards 1978).

The factors described above may variously limit our hearing ability, and systematic investigations are still needed with the census taker as the main study object. Studies on individuals as potential census takers should include objective examinations of the: (1) efficiency of bird identification at different levels of frequency and intensity, (2) pattern and speed of learning of bird songs in the ontogeny of a bird watcher, (3) number of song bouts needed for species identification, (4) parts of songs used as cues for identification, (5) effect of overlapping or masking on identification, (6) importance of the "out of range of birding" effect that occurs when an observer shifts from one locality to another, (7) effect of repetition on improvement of results, and (8) how these items vary in the application of different census methods. It will be important to formulate the problems carefully in order to compare the results with known check samples or parameters.

AN EXPERIMENT

Inasmuch as the effect of the human factor on census results is usually not correctly assessed, because of check sample bias, a test was designed to compare results obtained by different observers to a known sample. Observers (33) of varying quality, some of them currently involved in the Breeding Bird Survey, took part in the experiment. The aim was to examine the ability of these observers to discriminate sequences of species, species singing simultaneously and species from outside of the usual birding area of the observers. A total of 33 utterances from 12 species, arranged on a tape and delivered at intervals slightly longer than those heard at dawn hours, was played (Fig. 1). The observers had to identify the species and report them in sequence on a special checklist containing 36 species. The experiment was run twice.

The results of only 18 observers were kept for the analysis, because some observers did not complete the whole test during one or both runs or some have proved to be far from competent, identifying less than 20% of the birds. Table 1 shows how unrealistic were some estimates of bird numbers and comparisons between observers. Eighteen observers identified up to 27 species when only 12 were on tape, leading to discrepancies of up to 225% for number of species and 265% for number of individuals. The

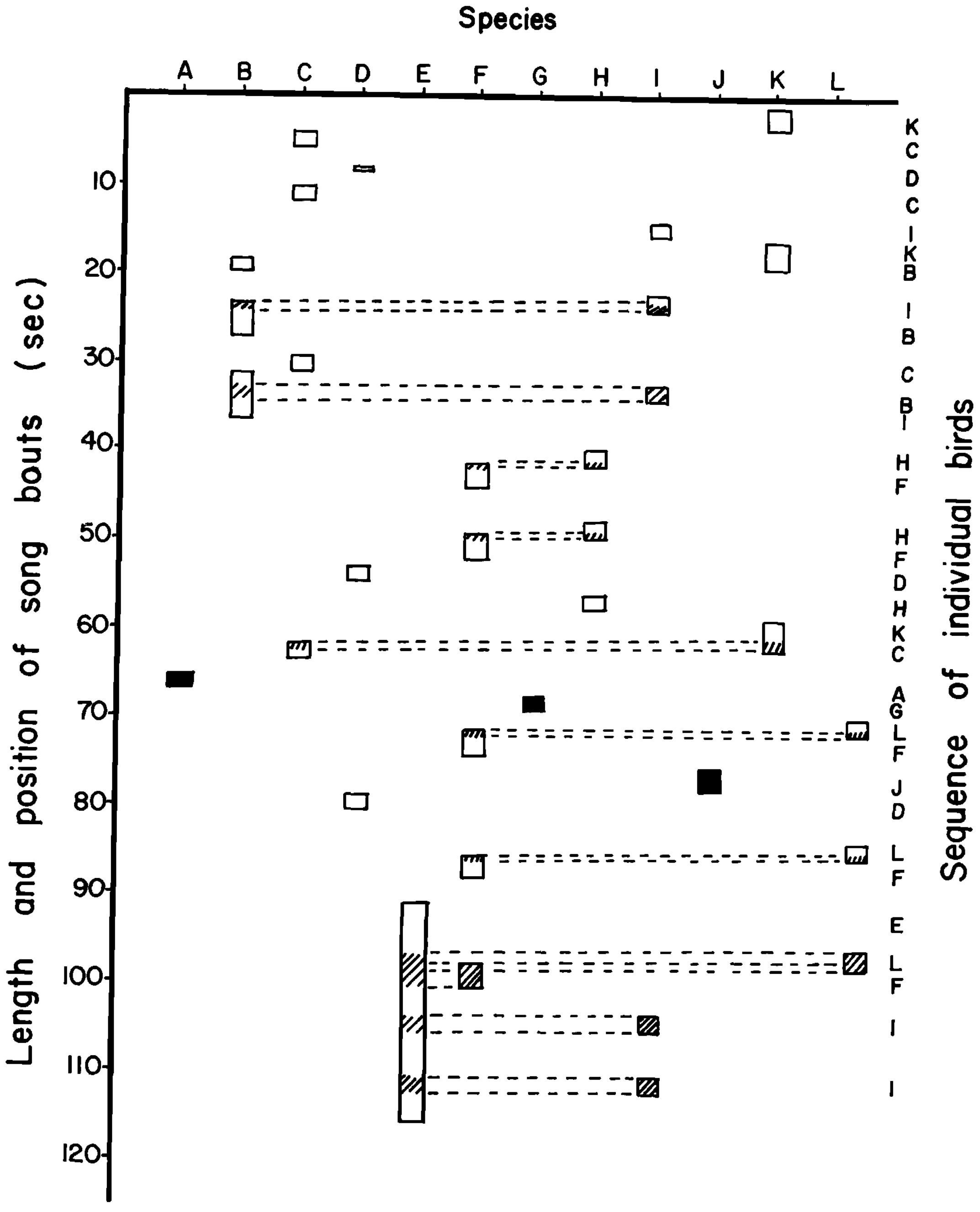

FIGURE 1. Arrangement of song bouts of 12 species on the tape recording used for the tests. Letters refer to the following species: A—*Nuttallornis borealis*; B—*Turdus migratorius*; C—*Catharus ustulatus*; D—*Catharus fuscescens*; E—*Vireo olivaceus*; F—*Vermivora peregrina*; G—*Dendroica virens*; H—*Dendroica pensylvanica*; I—*Wilsonia canadensis*; J—*Pinicola enucleator*; K—*Zonotrichia albicollis*; L—*Passerella iliaca*. Overlapping of songs is represented by broken lines and cross-hatching; black boxes are species singing only once on the recording.

TABLE 1
RESULTS OF ESTIMATES OF THE TAPE RECORDING CONTENT WHEN USING THE MAXIMUM, OR THE MEAN NUMBER OF BIRDS, OR THE RESULTS OBTAINED BY THE SUPPOSED BEST OBSERVER

	Number of birds						
Species	Max		Mean		Best		Tape
Nuttallornis borealis	2		0, 9		1		1
Turdus migratorius	5		1, 8		3		3
Catharus guttatus	4		1, 3		9		0
Catharus ustulatus	5		2, 8		4		4
Vermivora peregrina	4		0, 8		2		5
Dendroica petechia	3		0, 2		0		0
Dendroica magnolia	3		0, 2		3		0
Dendroica virens	1		0, 7		1		1
Dendroica pensylvanica	3		1, 5		3		3
Geothlypis trichas	2		0, 1		2		0
Wilsonia canadensis	6		2, 7		3		5
Pheucticus ludovicianus	3		0, 7		0		0
Pinicola enucleator	0		0		0		1
Pooecetes gramineus	3		0, 7		3		0
Zonotrichia albicollis	5		3, 2		3		3
Passerella iliaca	3		1, 2		0		3
Totals including errors							
Number of species[a]	27		27		14		12
Ratio over tape (%)		225		225		117	
Number of individuals (total)[a]	77		24, 3		27		29
Ratio over tape (%)		265		84		93	
Totals excluding errors							
Number of species[a]	11		11		10		12
Ratio over tape (%)		—		—		83	
Number of individuals (real)[a]	41		10, 1		24		29
Ratio over tape (%)		141		66		83	

[a] Species listing is incomplete, hence the discrepancy with the totals.

best overall estimate of the tape content was achieved with the results of the supposed best observer rather than the maximum or mean number of birds. But a 34% difference still occurs between the estimates by the best observer of species numbers with errors included and with errors excluded. This means that errors present partially cancel each other in estimates from census results. It is apparent that an unknown species may easily be unnoticed; for example, the summer song of the Pine Grosbeak (*Pinicola enucleator*) was unfamiliar to most of the observers, and it was unnoticed except by one observer in the first run. Some species, such as the Canada Warbler (*Wilsonia canadensis*) seem poorly known, being confused with 11 other species of birds. Confusion in counting the birds is also shown in the table.

Figure 2 shows that the number of individual birds correctly identified even on the second run tended to be directly correlated to the admitted use of hearing for bird identification by the observer (Spearman's $r = 0.52$, $P < 0.05$) and perhaps to the rating of the ability of the observer for the same purpose ($r = 0.40$, not significant). This means that using hearing more frequently increases the efficiency at identifying, as does practice. Repetition did improve the hearing and efficiency (Fig. 3). Although there seems to have been an improvement, the difference in the results between the runs was not significant ($\chi^2_{(11)} = 11.01$).

The masking phenomenon occurred in six overlapping singing situations. For example, the three Chestnut-sided Warblers (*Dendroica pensylvanica*) were correctly identified 3, 3, and 8 times in the first run and 5, 9, and 11 times out of 18 in the second. The first two songs of this warbler partly overlapped with one of a Tennessee Warbler (*Vermivora peregrina*), the last one did not overlap at all. The conclusion is self evident. Some of these conclusions do not pertain to hearing ability, but hearing ability is probably also correlated with species knowledge and training, due to the selectivity of response of the observers to the environmental stimuli (Lewis and Gower 1980). In fact, some observers mentioned not having heard the song of some of the species played on the tape!

In another study, we looked at the effect of different census methods on hearing ability. In this study, one observer censused the birds

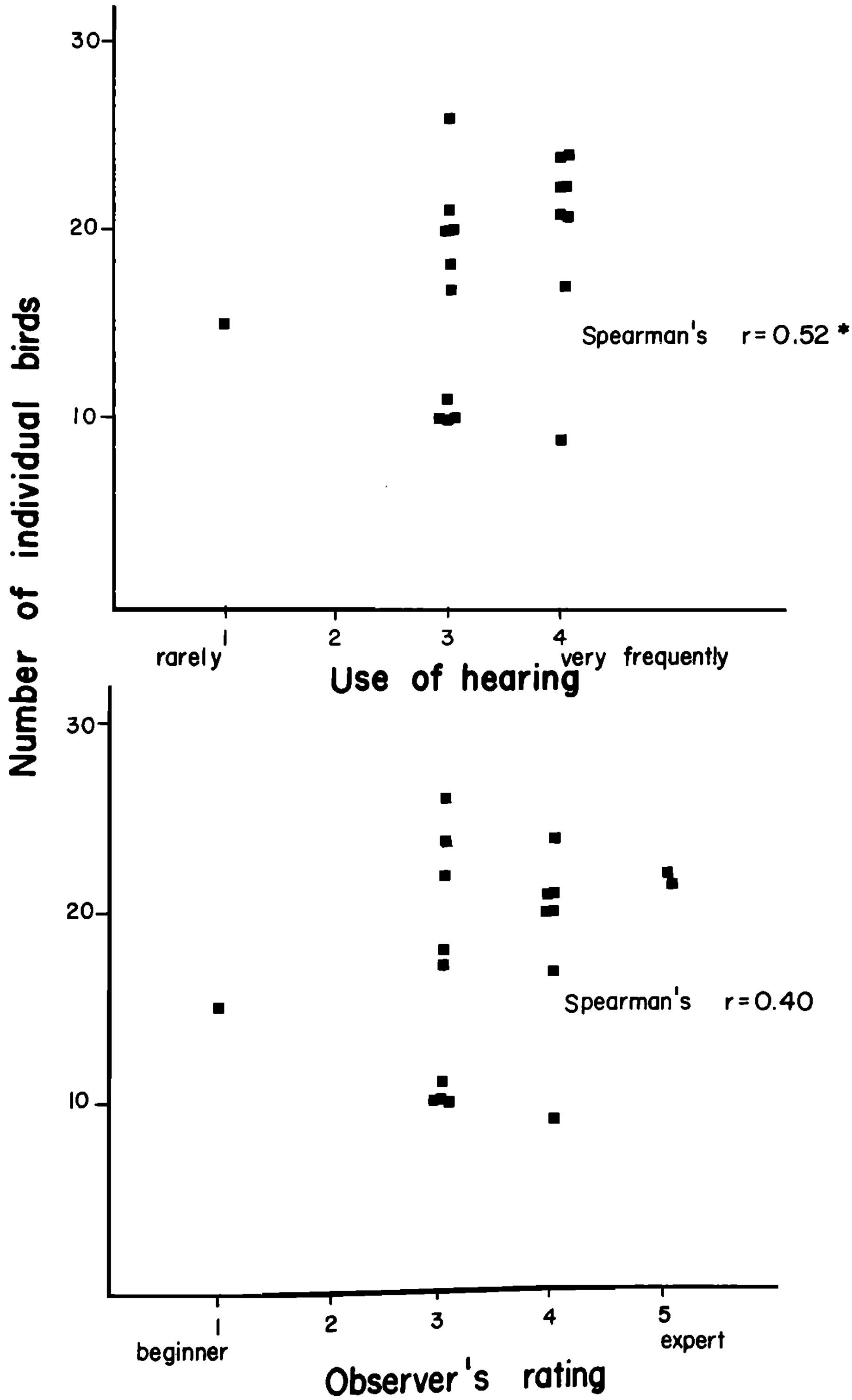

FIGURE 2. Relationship between the frequency of use of hearing by the observer or the observer's rating for field identification and the number of correctly identified individuals of birds over a possible maximum of 33 played on a tape.

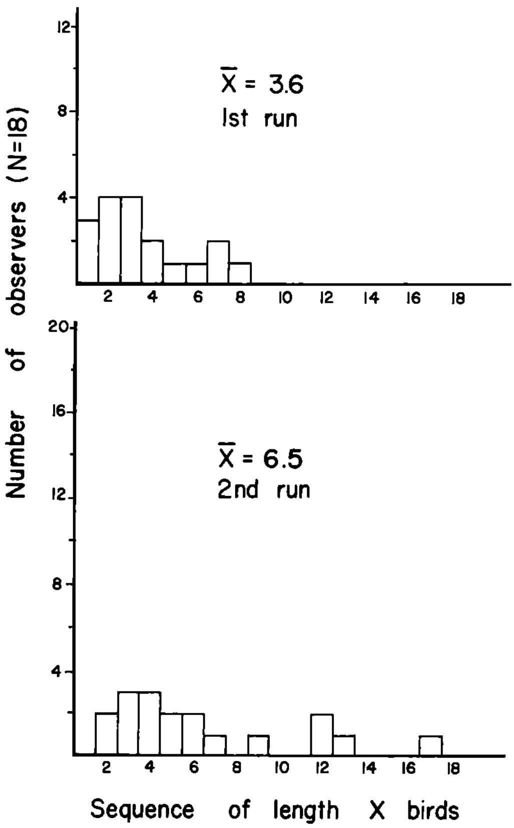

FIGURE 3. Effect of repetition on results of perception and identification of bird species. A sequence of, for example four, means that four individual adjacent birds were correctly identified along the played recording of 33 birds. The graphs show for each run of the test the number of observers that could identify correctly a maximum of χ birds in a sequence.

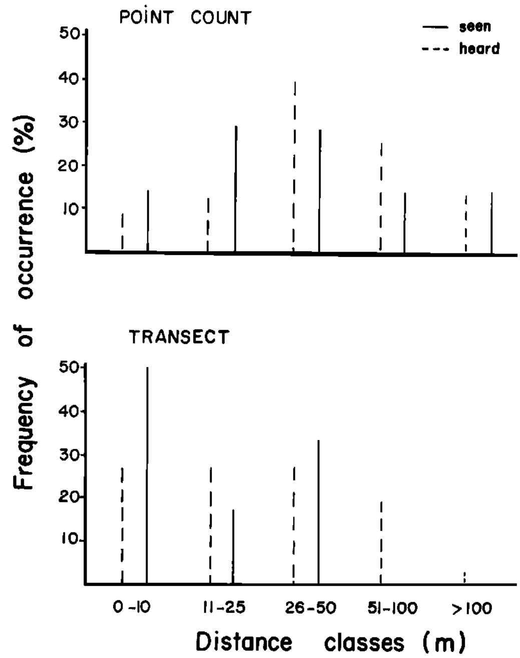

FIGURE 4. Importance of hearing in censusing and effect of a census method on hearing. The vertical bars represent the frequency of occurrence of birds (all species combined) heard or seen at different distances from the observer.

along a 3 km wooded path. The IPA method (point count method of Blondel et al. 1970) and the transect method were both used each week, but not simultaneously. The results of 16 weekly censuses from October to February were combined. Figure 4 shows that when the observer was walking he could not hear birds as far away as in point counts. On the other hand, standing for 10 minutes probably affected the activity of the birds near the observer. This figure also shows the importance of hearing in general, especially for detecting birds farther away from the observer. The relation between listening and looking would be much different in a breeding census situation. Further questions should be formulated to overcome hearing problems associated with different census methods.

DISCUSSION

Most researchers agree that observers do affect census results, whatever the census method used (Palmgren 1930, Enemar 1959). Examples of the qualities that could affect the census results are acuity of hearing, attentiveness, sensitivity in detecting individual birds, behavior of the observer on the terrain, emotional state, and others (Enemar 1962, Snow 1965, J. T. Emlen 1971, Best 1975, Berthold 1976, Enemar et al. 1978). For results obtained by the mapping technique, the error is often believed to be around 10% (see references in Berthold 1976), but preliminary evidence suggests that the error levels are probably higher than is usually and conveniently believed. Unfortunately, among the few studies involving comparisons of results obtained by different observers (Taylor 1965, Snow 1965, Enemar and Sjöstrand 1967, Enemar et al. 1978), only a few studies have made comparisons between results obtained simultaneously, or almost simultaneously (Carney and Petrides 1957; Enemar 1962, 1964; Hogstag 1967; Oelke et al. 1970; Jensen 1972). Unfortunately all of these studies except the one of Jensen compare the mean or maximum number of birds, or the results obtained by the supposed best observer for that plot. Studies are needed to evaluate the variability, the range of efficiency and the real impact of hearing ability on census results.

Audiology measurements do show differences from one individual to another, but most studies with humans are performed in relation to deafness (Martin 1975). To analyze this question more deeply, one would have to analyze observers from a psychological viewpoint to find out the reasons for the efficiency and performance of the different observers in species identification and in censusing. How, for example, can we explain the varying degrees of attentiveness of observers? How can we explain the different efficiencies in taking correct field notes? The lack of field or laboratory experiments on observers prevents further discussion.

We sought to answer the question: Are our perceptual tools able to analyze and decode properly the information sent by birds? We conclude that census takers need to improve the quality of their tools. It remains partly unknown how efficient we are at interpreting the emitted messages, in order to minimize the extent of the errors in census results. The extent to which we succeed at doing so needs to be reevaluated. Comparisons should be made with known check samples or parameters. Example of experiments would be to compare the results of observations with and without those obtained with a multimicrophone (multi-directionality) and a highly efficient recording device. It would be worthwhile to design tests in order to simulate the three dimensions for space locatability of bird songs played from a multiband recording device. Binaural hearing cues could be tested this way too. Other tests could simply play back especially arranged bird song sequences and ask the observer to identify species, or cues. Alterations of the songs could help to answer some of the above questions.

An increase in the number of cues used to perceive and identify the birds is definitely needed. One possibility is a wider use of sonograms, which have been overlooked in spite of their potential in the learning process, at least for those people who learn more visually than acoustically (Keith 1967, Beaver 1976). As in many other situations, progress depends on how one approaches the problem. Should we not also consider this strange creature, the census taker, as an object of scientific investigation!

ACKNOWLEDGMENTS

I wish to thank Ralph J. Raitt, J. Michael Scott, Patrick Bourgeron, C. John Ralph, and Jean-Marie Bergeron for generously criticizing different drafts of my manuscript. Thanks also to Nicole Vanasse who diligently retyped all the texts and Louis-Marie Thériault who made the drawings.